Amendments to the Specification:

Please replace the paragraph beginning at page 1, line 16, with the following amended paragraph:

--U.S. ~~application~~ Patent no. ~~08/796,792~~ 6,087,163 is incorporated herein by reference in it entirety.--

Please replace the paragraph beginning at page 13, line 21, with the following amended paragraph:

--The software used to manipulate and analyze protein sequences was available from public web servers or was part of the Genetics Computer Group (GCG) package [Wisconsin Package Version 9.1, Genetics Computer Group (GCG), Madison, Wisc.]. Customized C-Shell scripts were used to automate some of the tasks or to extract selected information from the output of some of the programs. Signal peptides were predicted with SPSCAN, which is part of the GCG package, and SignalP, a program originating from the Center for Biological Sequence Analysis at the Technical University of Denmark, Lyngby, Denmark and currently available on the Internet ~~at http://www.cbs.dtu.dk/services/SignalP~~. Putative transmembrane segments were identified with the program TMpred and prokaryotic membrane lipoprotein lipid attachment sites with the program PrositeScan, both programs originating from the Bioinformatics Group at the Swiss Institute for Experimental Cancer Research in Epalinges, Switzerland, and currently available on the Internet ~~at http://www.isrec.isb-sib.ch/software/TMPRED_form.html and http://www.isrec.isb-sib.ch/software/PSTSCAN_form.html, respectively~~. Protein similarity and relatedness was established with GAP and PILEUP, both in the GCG package, Blast originating from the National Center for Biotechnology Information of the National Institutes for Health, Bethesda, MD ~~and currently available on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/~~, and AllAll originating from the Swiss Institute of Technology, Zurich, Switzerland, and currently available on the Internet ~~at http:cbrg.inf.ethz.ch/subsection3_1_1.html~~.--

Please replace the paragraph beginning at page 14, line 16, with the following amended paragraph:

--The amino acid sequences of the 3924 proteins predicted by the analysis of the *M. tuberculosis* genomic sequence have been made available by the Sanger Centre, Cambridge, England, and were downloaded from the current Sanger Center website ~~[http://www.sanger.ac.uk/Projects/M_tuberculosis/]~~. Segments containing the first 70 amino acids of each predicted protein were analyzed by a system of our own design utilizing two different computer programs (SPSCAN and SignalP) designed to predict the occurrence of signal peptides. We concluded that combining the output from the two programs would increase the reliability of the selection. Both programs can detect signal peptides in polypeptides from eukaryotic and prokaryotic organisms, including gram-positive and gram-negative bacteria. To analyze the *M. tuberculosis* proteins the gram-positive mode was used. We performed an analysis with SPSCAN allowing only one prediction per protein, setting the minimum score threshold at -10, both in the standard and the adjusted modes. In the adjusted mode, signal peptides longer than a certain threshold value are penalized. We found that the correlation between the scores obtained with SPSCAN in the standard and adjusted modes increased with the value of the score, i.e., signal peptides that received high scores in standard mode also had high scores in the adjusted mode. We determined to use only the adjusted mode in subsequent steps.--

Please replace the title of Table 4 on page 20, line 1, with the following amended title:

--Table 4. Presence <u>of</u> mtsp coding regions in various strains of *Mycobacterium tuberculosis*.--